Discrete Applied Mathematics 9 (1984) 69–75
North-Holland

A NOTE ON OFF-LINE MACHINES WITH 'BROWNIAN' INPUT HEADS

Keijo RUOHONEN
Institute of Mathematics, Tampere University of Technology, SF-33101 Tampere 10, Finland

Received 9 June 1982
Revised 14 September 1983

An off-line machine is a Turing machine with a separate read-only input tape. Such a machine is said to have a 'Brownian' input head if it has no control over and cannot observe the moves of the read-only head scanning the input tape. It is shown in this note that the binary languages (i.e., languages over a two-symbol alphabet) recognized by such machines are all regular. The proof is not constructive and relies heavily on the so-called 'finite sequences theorem' in the theory of well-partial-ordering.

1. Introduction and preliminaries

For a so-called *off-line machine* the input is given on a separate input tape between endmarkers. The head scanning the input tape is of the read-only type and it cannot pass the endmarkers. The rest of the machine consists of a central unit with a finite memory, a number of work tapes (possibly multidimensional), oracles, etc. The moves of the head scanning the input tape are *controlled* by the central unit according to the computations performed on the work tapes, the 'advice' obtained from the oracles, etc. These moves are also *observed* by the central unit in the sense that the unit 'knows' which move the head takes.

In the present note we are interested in what happens if we remove both the controllability and the observability of the input head. To define precisely what this means, let us consider an off-line machine $M=(A,\lambda,S)$ where A is the input alphabet, λ is the transfer function and S specifies the remaining part of the machine. The mode of acceptance should be defined, too, but it can be rather arbitrary (halting, nonhalting, etc.). Given a pair (x,s), where x is the input symbol or endmarker under scan on the input tape and s contains all other data needed for the transfer, $\lambda(x,s)$ is a set (possibly empty) of pairs of the form (t,i), where i indicates the move of the input head ("+1" = "move to the right", "−1" = "move to the left" and "0" = "no move") and t contains all other data necessary for specification of the next configuration of the machine.

We say that M has a *Brownian input head* if, for all s and t,

$$\text{(I)}\quad (t,+1)\in\lambda(x,s) \Leftrightarrow (t,-1)\in\lambda(x,s) \Leftrightarrow (t,0)\in\lambda(x,s), \quad \text{when } x\in A,$$

$$\text{(II)}\quad (t,+1)\in\lambda(x,s) \Leftrightarrow (t,0)\in\lambda(x,s),\quad \text{when } x \text{ is the left endmarker,}$$

$$\text{(III)}\quad (t,-1)\in\lambda(x,s) \Leftrightarrow (t,0)\in\lambda(x,s),\quad \text{when } x \text{ is the right endmarker.}$$

This means that M has no control over the moves of its input head. It also means that M cannot observe the moves of its input head. (The corresponding definition of noncontrollability without nonobservability, corresponding to (I) above, is the following: for all s and all $x\in A$,

$$(\exists t_1)\ [(t_1,+1)\in\lambda(x,s)] \Leftrightarrow (\exists t_2)\ [(t_2,-1)\in\lambda(x,s)]$$
$$\Leftrightarrow (\exists t_3)\ [(t_3,0)\in\lambda(x,s)].)$$

Such an input head can be considered as performing a random walk or a discrete bounded Brownian movement on the input tape. From now on we will consider only off-line machines with Brownian input heads (*Brownian off-line machines*, in short).

If the input alphabet A is unary, say $A=\{a\}$, then the (unary) languages recognized by Brownian off-line machines are all regular, regardless of the underlying structure of the machine or the mode of acceptance. Indeed, such a language is always either empty or of one of the forms a^*, a^+, $\{a,a^2,\dots,a^n\}$, $\{\Lambda,a,a^2,\dots,a^n\}$, as is easily seen ($\Lambda$ denotes the empty word). All these languages are, in fact, recognized already by Brownian finite automata (i.e., Brownian off-line machines with the central unit as their only structure). Note that if the underlying structure of these Brownian off-line machines is allowed to be that of general Turing machines, then specifying the form of the language is ineffective because emptiness of unary recursively enumerable languages is undecidable.

On the other hand, if the input alphabet is ternary, say $A=\{a,b,c\}$, then, assuming the halting mode of acceptance, we get essentially everything recursively enumerable: given any recursively enumerable language L over a unary alphabet, say $\{a\}$, the language $\{(abc)^n \mid a^n\in L\}$ over A equals the language recognized by a Brownian off-line machine after intersecting with the regular language $(abc)^*$. The idea is that, for inputs in $(abc)^*$, changing or nonchanging of the input symbol under scan of the input head gives an 'artificial observability' of the motions of the head which suffices for a nondeterministic simulation of a non-Brownian off-line machine recognizing L.

This situation renders the case of binary input alphabets interesting. We will prove the somewhat surprising result that the binary languages recognized by Brownian off-line machines are all regular and are, in fact, recognized already by Brownian finite automata. This result seems to be rather less evident than the corresponding result for unary input alphabets. Our main tool is the so-called 'finite sequences theorem' in the theory of well-partial-ordering, see e.g. [3].

A reader not sufficiently familiar with elements of automata theory is referred to [2].

2. The result

We will make use of a partial order of the input words of our Brownian off-line machines. We fix the endmarkers as follows: ¢ is the left endmarker and § is the right endmarker. To define the partial order, we need the concept of a random input sequence generated by an input word $P=x_1x_2\cdots x_m$ where $x_1,x_2,\ldots,x_m \in A$. For simplicity we denote the left endmarker ¢ also by x_0 and the right endmarker § by x_{m+1}. A sequence $a_0,a_1,\ldots$ of symbols of $A \cup \{¢, §\}$ (finite or infinite) is said to be a *random input sequence generated by P* if there exists a sequence $i_0,i_1,\ldots$ of indices from the set $\{0,1,\ldots,m+1\}$ such that

(i) $a_n = x_{i_n}$ for $n = 0,1,\ldots,$

(ii) $|i_n - i_{n+1}| \le 1$ for $n=0,1,\ldots,$ and

(iii) $a_0 = ¢$, i.e. $i_0 = 0$.

Let us denote by C_P the set of all random input sequences generated by P. We then define the binary relation $\propto$ on A^* by

$$P \propto Q \Leftrightarrow C_Q \subset C_P.$$

Obviously $\propto$ is reflexive and transitive. Since, for a word $Q=x_1x_2\cdots x_m$, where $x_1,x_2,\ldots,x_m \in A$, the sequence $¢,x_1,x_2,\ldots,x_m,§$ is in C_Q, we know that

$$P \propto Q \quad \text{implies} \quad |P| \le |Q|$$

($|X|$ denotes the length of the word X) and

$$P \propto Q \text{ and } |P| = |Q| \quad \text{imply} \quad P = Q.$$

Thus $\propto$ is also antisymmetric and hence a partial order on A^*.

The following lemma is frequently used.

Lemma 1. *Let P and Q be words over A, and $Q=x_1x_2\cdots x_m$ where $x_1,x_2,\ldots,x_m \in A$. Then $P\propto Q$ if and only if $¢,x_1,x_2,\ldots,x_m,§$ is in C_P.*

Proof. Clearly $¢,x_1,x_2,\ldots,x_m,§$ is in C_Q and hence also in C_P if $P\propto Q$. Assume then that $¢,x_1,x_2,\ldots,x_m,§$ is in C_P and let $i_0,i_1,\ldots,i_{m+1}$ be the corresponding index sequence. Suppose $a_0,a_1,\ldots$ is a random input sequence generated by Q with the index sequence $i_0',i_1',\ldots,$ and define $i_j''=i_{i_j'}$ for $j=0,1,\ldots$. It is then easily verified that $i_0'',i_1'',\ldots$ is an index sequence corresponding to generation of $a_0,a_1,\ldots$ by P. We deduce that $C_Q \subset C_P$ whence $P\propto Q$. □

Our next task is to show that $\propto$ is a well-partial-order if A is binary, say $A=\{a,b\}$. Recall that a well-partial-order is a partial order such that

(1) every strictly descending sequence is finite, and

(2) every set of pairwise incomparable elements is finite.

For this purpose, let us define first the set T of triples as follows: $T = \{a,b\} \times N_+ \times \{I, F, M, IF\}$ (N_+ denotes the set of positive integers, and "I" indicates an initial block, "F" a final block, "M" a middle block and "IF" a block with both initial and final). An order on T is given by

$$(x_1, n_1, Y_1) \leq_1 (x_2, n_2, Y_2) \Leftrightarrow X_1 = X_2 \;\&\; n_1 \leq n_2 \;\&\; Y_1 = Y_2.$$

Obviously T is well-partially-ordered by $\leq_1$.

Now each word P over $\{a,b\}$ can be interpreted as a sequence U_P of elements of T in the following way:

(A) U_Λ is the empty sequence.

(B) If $P = x^n$, where $x \in \{a,b\}$ and $n > 0$, then U_P is the singleton sequence (x, n, IF).

(C) If $P = x^{n_1}y^{m_1}x^{n_2}y^{m_2}\cdots x^{n_k}y^{m_k}$, where $\{x,y\} = \{a,b\}$ and $n_1, m_1, \ldots, n_k, m_k > 0$, then U_P is $(x, n_1, I), (y, m_1, M), (x, n_2, M), (y, m_2, M), \ldots, (x, n_k, M), (y, m_k, F)$.

(D) If $P = x^{n_1}y^{m_1}x^{n_2}y^{m_2}\cdots x^{n_k}y^{m_k}x^{n_{k+1}}$, where $\{x,y\} = \{a,b\}$ and $n_1, m_1, \ldots, n_k, m_k, n_{k+1} > 0$, then U_P is $(x, n_1, I), (y, m_1, M), (x, n_2, M), (y, m_2, M), \ldots, (x, n_k, M), (y, m_k, M), (x, n_{k+1}, F)$.

Let us then define the binary relation $\leq_2$ on $\{U_P \mid P \in \{a,b\}^*\}$ by

$$U_P \leq_2 U_Q \Leftrightarrow \text{either } P = Q = \Lambda \text{ or } P \neq \Lambda \text{ and } U_P \text{ is termwise majorated by a subsequence of } U_Q \text{ under } \leq_1.$$

The so-called 'finite sequences theorem' in the theory of well-partial-ordering (see e.g. [3]) then immediately implies that $\leq_2$ is well-partial-order. On the other hand, we have

Lemma 2. *$P \propto Q$ if and only if $U_P \leq_2 U_Q$.*

Proof. We define first a binary relation ϱ on $\{a,b\}^*$ and then obtain $\propto$ as the reflexive-transitive closure of ϱ. We list all situations where $P \varrho Q$ holds true:

(a) $P = P_1 a P_2$ and $Q = P_1 a^2 P_2$,

(b) $P = P_1 b P_2$ and $Q = P_1 b^2 P_2$,

(c) $P = P_1 ab P_2$ and $Q = P_1 (ab)^2 P_2$,

(d) $P = P_1 ba P_2$ and $Q = P_1 (ba)^2 P_2$.

To show that $\varrho^* = \propto$ we assume first that $P \varrho^* Q$ and show that $P \propto Q$ (ϱ^* denotes the reflexive-transitive closure of ϱ). Since $\propto$ is reflexive and transitive, it suffices to consider the case where $P \varrho Q$. By Lemma 1, it also suffices to show that, if $Q = x_1 x_2 \cdots x_m$ where $x_1, x_2, \ldots, x_m \in \{a,b\}$, then $¢, x_1, x_2, \ldots, x_m, \S$ is a random input sequence generated by P, but this is immediate from the definition of ϱ.

Assume then that $P \propto Q$. Since this implies $|P| \leq |Q|$, it follows that there are only finitely many words R such that $P \propto R \propto Q$. By the reflexiveness and transitiveness of ϱ^*, it then suffices to consider the case where $P \neq Q$ and $P \propto R \propto Q$ implies $R = P$

or $R=Q$, and show that in this case $P\varrho Q$. Let $P=x_1x_2\cdots x_m$ and $Q=a_1a_2\cdots a_r$ where $x_1,x_2,\ldots,x_m,a_1,a_2,\ldots,a_r\in A$. By Lemma 1 we then know that $¢,a_1,a_2,\ldots,a_r,\S$ is a random input sequence generated by P. Let $i_0,i_1,\ldots,i_{r+1}$ be the corresponding sequence of indices from $\{0,1,\ldots,m+1\}$, where $i_0=0$ and $i_{r+1}=m+1$ and $1\le i_n\le m$ for all n, $1\le n\le r$.

Suppose first that $i_n=i_{n+1}$ for a value of n, $1\le n\le r$. Then $i_0,i_1,\ldots,i_n,i_{n+2},\ldots,i_{r+1}$ is also a legitimate index sequence giving the random input sequence $¢,a_1,a_2,\ldots,a_n$, $a_{n+2},\ldots,a_r,\S$ generated by P. By Lemma 1, then $P\propto R\propto Q$ for $R=a_1a_2\cdots a_na_{n+2}\cdots a_r$ whence $R=P$ and either (a) or (b) follows.

Suppose then that $i_n\ne i_{n+1}$ for all n, $1\le n<r$. Since $P\propto Q$ and $P\ne Q$, we have $i_{n-1}=i_{n+1}=i_n-1$ for a value of n, $1<n<r$. If now $x_{i_n}=x_{i_{n+1}}$, the legitimate index sequence $i_0,i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_{r+1}$ gives the word $R=a_1a_2\cdots a_{n-1}a_{n+1}\cdots a_r$ for which $P\propto R\propto Q$, and we must have $R=P$ whence again either (a) or (b) holds. So let $x_{i_n}\ne x_{i_{n+1}}$. Denote by k the smallest number such that $k>n+1$ and $i_k=i_n$ (note that such a k exists). The sequence $i_0,i_1,\ldots,i_{n-1},i_k,\ldots,i_{r+1}$, as well as the sequence $i_0,i_1,\ldots,i_{n+1},i_k,\ldots,i_{r+1}$, is a legitimate index sequence. By Lemma 1, $P\propto R_1\propto R_2$ where $R_1=a_1a_2\cdots a_{n-1}a_k\cdots a_r$ and $R_2=a_1a_2\cdots a_{n+1}a_k\cdots a_r$. Again using Lemma 1, we see easily that $a_na_{n+1}\propto a_na_{n+1}\cdots a_{k-1}$ whence also $R_2\propto Q$. (Recall that $a_n\ne a_{n+1}$ and note that $i_{k-1}=i_{n+1}$ by the minimality of k whence $a_{k-1}=a_{n+1}$.) It follows that $R_1=P$ and $R_2=Q$ and so either (c) or (d) holds true.

We can thus conclude that $\varrho^*=\propto$. It remains to be shown that

$$P\varrho^*Q \Leftrightarrow U_P\le_2 U_Q.$$

Since $\le_2$ is reflexive and transitive, it suffices to show the implication from left to right in the case where $P\varrho Q$, and this is immediate.

Since $U_P\le_2 U_Q$ implies $|P|\le|Q|$, there are only finitely many words R such that $U_P\le_2 U_R\le_2 U_Q$. By the reflexiveness and transitiveness of ϱ^*, it then suffices to show the implication from right to left in the case where $U_P\ne U_Q$ (i.e., $P\ne Q$), and $U_P\le_2 U_R\le_2 U_Q$ implies $R=P$ or $R=Q$. But in this case it is easily verified that we necessarily have $P\varrho Q$. □

Thus $\propto$ *is a well-partial-order,* too. It is a property of a well-partially-ordered set that every nonempty subset of it has at least one but no more than a finite number of minimal elements. In analogy to Theorem 3 of [1] we have

Lemma 3. *Let L be a nonempty language over $\{a,b\}$ such that*

$$P\in L \text{ and } P\propto Q \quad \text{imply} \quad Q\in L.$$

Then L is regular.

Proof. Let F be the set of minimal elements of L. Thus F is a nonempty finite set and

$$L=\{Q\mid P\propto Q \text{ and } P\in F\}.$$

The regularity of L is then easily seen by a finite automaton construction based on Lemma 1 or a regular expression construction based on the characterization $\propto = \varrho^*$ given in the proof of Lemma 2. □

We are now ready for our main result.

Theorem 1. *For any Brownian off-line machine M with a binary input alphabet, the language recognized by M is regular.*

Proof. Assume that Q is accepted by M and $P \propto Q$. During the computation (halting or nonhalting) leading to the acceptance of Q by M, the central unit receives a certain random input sequence (finite or infinite), generated by Q, from its Brownian input head. Since $C_Q \subset C_P$, M receives the same random input sequence with input P, too, and hence accepts P. We deduce that

$$P \in \bar{L} \text{ and } P \propto Q \quad \text{imply} \quad Q \in \bar{L},$$

where $\bar{L}$ denotes the complement of L, and the theorem follows from Lemma 3 and the well-known closure of regular languages under complementation. □

Corollary 1. *For any Brownian off-line machine with a binary input alphabet there is an equivalent Brownian finite automaton.*

Proof. Let M be a Brownian off-line machine with input alphabet $\{a, b\}$. By Theorem 1, the language L recognized by M is regular and thus also recognized by a deterministic finite automaton M_1. We then transform M_1 to an equivalent Brownian finite automaton M_2 as follows. Let the set of states of M_1 be S_1 and let its initial state (resp. set of final states) be s_0 (resp. S_f). Then the set of states of M_2 is $S_2 = S_1 \cup \{\alpha, \beta\}$ where α is the (only) initial state and β is a 'nonhalting state'. Let the (total) transfer function of M_1 be $\lambda_1 : (S_1 \times \{a, b\}) \to S_1$. The transfer function $\lambda_2 : (S_2 \times \{a, b, ¢, \S\}) \to 2^{S_2 \times \{-1, 0, +1\}}$ of M_2 is then given by

$$\lambda_2(s, ¢) = \begin{cases} \{s_0\} \times \{0, +1\}, & \text{if } s = \alpha, \\ \{\beta\} \times \{0, +1\}, & \text{if } s \in S_2 - \{\alpha\}, \end{cases}$$

$$\lambda_2(s, x) = \begin{cases} \{\lambda_1(s, x)\} \times \{-1, 0, +1\}, & \text{if } s \in S_1 \text{ and } x \in \{a, b\}, \\ \{\beta\} \times \{-1, 0, +1\}, & \text{if } s \in S_2 - S_1 \text{ and } x \in \{a, b\}, \end{cases}$$

$$\lambda_2(s, \S) = \begin{cases} \text{empty}, & \text{if } s \in S_f, \\ \{\beta\} \times \{-1, 0\}, & \text{if } s \in S_2 - S_f. \end{cases}$$

The mode of acceptance of M_2 is halting. It is readily seen that any word of L is accepted by M_2. On the other hand, if a word P is accepted by M_2, then, by Lemma 1, L contains a word Q such that $P \propto Q$, whence P is in L, too. □

Acknowledgement

The author wishes to thank the anonymous referee for useful suggestions.

References

[1] L.H. Haines, On free monoids partially ordered by embedding, J. Combin. Theory 6 (1969) 94–98.

[2] J.E. Hopcroft and J.D. Ullman, Introduction to Automata Theory, Languages and Computation (Addison-Wesley, Reading, MA, 1979).

[3] J.B. Kruskal, The theory of well-quasi-ordering: a frequently discovered concept, J. Combin. Theory (A) 13 (1972) 297–305.